JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **121**, 487–505 (1987)

On a Multiplication and a Theory of Integration for Belief and Plausibility Functions*

MANFRED BERRES

Institut für Medizinische Statistik und Dokumentation, Johannes Gutenberg-Universität, Mainz, West Germany

Submitted by L. A. Zadeh

Received May 17, 1985

Belief and plausibility functions have been introduced as generalizations of probability measures, which abandon the axiom of additivity. It turns out that elementwise multiplication is a binary operation on the set of belief functions. If the set functions of the type considered here are defined on a locally compact and separable space X, a theorem by Choquet ensures that they can be represented by a probability measure on the space containing the closed subsets of X, the so-called basic probability assignment. This is basic for defining two new types of integrals. One of them may be used to measure the degree of non-additivity of the belief or plausibility function. The other one is a generalization of the Lebesgue integral. The latter is compared with Choquet's and Sugeno's integrals for non-additive set functions.

1. INTRODUCTION

The concept of belief function and plausibility function has been introduced and discussed by Shafer [8–10] to model systems which are not necessarily controlled by chance. Given a set X of possible outcomes and a subset $A \subset X$, the belief function $\mathrm{Bel}(A)$ measures the strength of evidence which supports A. There may be evidence which cannot discriminate between A and its complement A^c. Therefore $\mathrm{Bel}(A^c)$ may be less than $1 - \mathrm{Bel}(A)$. A belief function Bel induces a conjugate plausibility function Pl by $\mathrm{Pl}(A) = 1 - \mathrm{Bel}(A^c)$ which measures the extent to which one does not believe in the contrary of A.

Shafer [9] intended to present a generalization of subjective probabilities which is not necessarily additive. But Walley and Fine [12] describe several situations in which non-additivity is desirable, even from a frequentist point of view, e.g., if independent observations are governed by dif-

* This article is a part of author's Ph.D. thesis.

0022-247X/87 $3.00

ferent probability laws, and $\underline{\lim}$ and $\overline{\lim}$ have to be considered. In robust statistics other examples of non-additive set functions arise from contaminated probability measures (e.g., [5]). There is a close relationship between belief and plausibility functions on the one hand and Dempster's [3] lower and upper probabilities on the other hand. This has been described by Nguyen [7].

The present article is based on the fact that belief and plausibility functions can be represented by a probability measure m defined on the space of all closed subsets of X which we call basic probability assignment in accordance with Shafer [9]. For finite sets X the existence of m has been shown by Shafer [9]. The general case has been treated by Shafer [8, 10] and Matheron [6]. We prefer to use Matheron's approach, because it seems to be more natural and more closely related to Choquet's [2] paper on capacities which is fundamental for the mathematical theory of belief and plausibility functions.

Dempster [3] has defined a binary operation in order to combine independent sources of evidence which he calls orthogonal sum. In Section 3 we introduce a simpler binary operation, namely, the multiplication of belief functions. Its properties are different from those of Dempster's orthogonal sum.

In Sections 4 and 5 we define two types of integrals based on belief and plausibility functions instead of probabilities. The first one can be used to measure the extent of non-additivity of the set function itself, whereas the second one is a generalization of the Lebesgue integral, i.e., it can be interpreted as an expectation. The value of such an integral is the same for every conjugate pair of belief and plausibility functions by definition. This property is desirable, since a belief function and its conjugate plausibility function reflect the same attitude towards a partially known reality from different points of view.

We prove some limit theorems for these integrals, study several examples, and investigate the integrals' properties under the binary operations mentioned above. Finally, we compare the second integral with Choquet's [2] and Sugeno's [11] integrals for non-additive set functions.

2. Belief and Plausibility Functions

Consider a locally compact, separable (LCS) space X. Let $\mathscr{B}(X)$ be the σ-algebra of borel sets generated by the open sets in X. We denote by $\mathscr{F}=\mathscr{F}(X)$, $\mathscr{K}=\mathscr{K}(X)$, and $\mathscr{G}=\mathscr{G}(X)$ the classes of closed, compact, and open subsets of X, respectively. For arbitrary $B \subset X$ let

$$\mathscr{F}_B=\{F\in\mathscr{F}: F\cap B\neq\varnothing\} \qquad \text{and} \qquad \mathscr{F}^B=\{F\in\mathscr{F}: F\cap B=\varnothing\}.$$

The two families $\mathscr{F}^K$, $K\in\mathscr{K}$, and $\mathscr{F}_G$, $G\in\mathscr{G}$, generate a topology $\mathscr{T}(\mathscr{F})$ on

$\mathscr{F}$. The open sets of $\mathscr{T}(\mathscr{F})$ generate a σ-algebra of subsets of $\mathscr{F}$ which will be denoted by $\mathscr{S}(\mathscr{F})$.

A function $\mathrm{Pl}: \mathscr{B}(X) \to [0, 1]$ is called a *plausibility function*, if it satisfies

(i) $\mathrm{Pl}(\varnothing) = 0$, $\mathrm{Pl}(X) = 1$,

(ii) $A, B \in \mathscr{B}(X)$, $A \subset B \Rightarrow \mathrm{Pl}(A) \leqslant \mathrm{Pl}(B)$,

(iii) $A_n \in \mathscr{B}(X)$, $A_n \uparrow A \Rightarrow \mathrm{Pl}(A_n) \uparrow \mathrm{Pl}(A)$,

(iv) $K_n \in \mathscr{K}(X)$, $K_n \downarrow K \Rightarrow \mathrm{Pl}(K_n) \downarrow \mathrm{Pl}(K)$,

(v) Pl is *alternating of order* ∞, i.e.,

$$\mathrm{Pl}\left(\bigcap_{i=1}^{n} A_i\right) \leqslant \sum_{\varnothing \neq I \subset \{1,\ldots,n\}} (-1)^{|I|+1} \mathrm{Pl}\left(\bigcup_{i \in I} A_i\right)$$
$$\text{for all } n \geqslant 1;\ A_1, \ldots, A_n \in \mathscr{B}(X). \tag{2.1}$$

A function $\mathrm{Bel}: \mathscr{B}(X) \to [0, 1]$ is called a *belief function*, if the corresponding properties (i′)–(v′) hold with the following modifications:

(iii′) $A_n \in \mathscr{B}(X)$, $A_n \downarrow A \Rightarrow \mathrm{Bel}(A_n) \downarrow \mathrm{Bel}(A)$,

(iv′) $G_n \in \mathscr{G}(X)$, $G_n \uparrow G$, $G_1^c \in \mathscr{K}(X) \Rightarrow \mathrm{Bel}(G_n) \uparrow \mathrm{Bel}(G)$,

(v′) Bel is *monotone of order* ∞, i.e.,

$$\mathrm{Bel}\left(\bigcup_{i=1}^{n} A_i\right) \geqslant \sum_{\varnothing \neq I \subset \{1,\ldots,n\}} (-1)^{|I|+1} \mathrm{Bel}\left(\bigcap_{i \in I} A_i\right)$$
$$\text{for al } n \geqslant 1 \text{ and } A_1, \ldots, A_n \in \mathscr{B}(X). \tag{2.2}$$

It should be remarked that conditions (v) and (v′) imply subadditivity and superadditivity, respectively.

Now, the existence of a *basic probability assignment* (*b.p.a.*) m is guaranteed by the famous

THEOREM (Choquet [2], Matheron [6]). *A function* $\mathrm{Pl}: \mathscr{K}(X) \to [0, 1]$ *satisfying* $\sup_{K \in \mathscr{K}} \mathrm{Pl}(K) = 1$ *is a plausibility function, if and only if there exists a probability measure* m *on* $(\mathscr{F}, \mathscr{S}(\mathscr{F}))$ *such that* $\mathrm{Pl}(K) = m(\mathscr{F}_K) \cdot m$ *is uniquely determined.*

As Matheron [6, p. 29f.] pointed out, Pl can be extended from $\mathscr{K}(X)$ to $\mathscr{B}(X)$ (and even to the power set of X), and $\mathscr{F}_B$ is measurable with respect to the completion $(\mathscr{F}, \tilde{\mathscr{S}}(\mathscr{F}), \tilde{m})$ of $(\mathscr{F}, \mathscr{S}(\mathscr{F}), m)$ for every $B \in \mathscr{B}(X)$; hence the equality $\mathrm{Pl}(B) = \tilde{m}(\mathscr{F}_B)$ is well-defined for all $B \in \mathscr{B}(X)$. For simplicity of notation we will omit the tilde.

Pl is a plausibility function, if and only if $\mathrm{Bel}(B) = 1 - \mathrm{Pl}(B^c)$ $(B \in \mathscr{B}(X))$ is a belief function. On the other hand, if a belief function Bel and a plausibility function Pl satisfy the previous equation, then we call them *conjugate* to each other. Thus a b.p.a. m simultaneously generates the conjugate functions $\mathrm{Pl}(B)=m(\mathscr{F}_B)$ and $\mathrm{Bel}(B)=m(\mathscr{F}^{B^c})$, and every belief function can be represented by a b.p.a., too.

Clearly, $\mathrm{Pl}(B)=m(\{F\in\mathscr{F}\mid F\cap B\neq\emptyset\})$ is the probability of the maximal subset of $\mathscr{F}$ which *can attribute* to B, and $\mathrm{Bel}(B)=m(\{F\in\mathscr{F}\mid F\subset B\})$ is the probability of the minimal subset of $\mathscr{F}$ which *attributes* to B. That is why Pl and Bel are sometimes called *upper and lower probabilities* (Dempster [3]).

EXAMPLE 1. A probability measure P on $(X, \mathscr{B}(X))$ is alternating and monotone of order ∞, as equality holds true in (2.1) and (2.2) by the Sylvester–Poincaré equality. Now $P(B)=m(\mathscr{F}_B)=m(\mathscr{F}^{B^c})$ and therefore

$$m(\mathscr{F}_B\backslash\mathscr{F}^{B^c})=m(\{F\in\mathscr{F}\mid F\cap B\neq\emptyset \text{ and } F\cap B^c\neq\emptyset\})=0.$$

If X is finite and discrete, and $C\subset X$ contains more than one element, this implies that $m(\{C\})=0$. Thus m is concentrated on singletons.

EXAMPLE 2. Let P be a probability measure on $(X, \mathscr{B}(X))$ and $\varepsilon\in[0,1]$.

It is easily verified that $\mathrm{Bel}(B)=(1-\varepsilon)\,P(B)$, $B\neq X$, and $\mathrm{Bel}(X)=1$ define a belief function. Its conjugate plausibility function Pl given by $\mathrm{Pl}(B)=(1-\varepsilon)\,P(B)+\varepsilon$ for $B\neq\emptyset$ and $\mathrm{Pl}(\emptyset)=0$ was called *ε-contamination* by Huber and Strassen [5]. Its b.p.a. takes the values

$$m(\mathscr{F}_B)=(1-\varepsilon)\,P(B)+\varepsilon \quad\text{and}\quad m(\mathscr{F}^{B^c})=(1-\varepsilon)\,P(B) \quad\text{for}\quad B\neq\emptyset, X,$$

so $m(\mathscr{F}_B\backslash\mathscr{F}^{B^c})=\varepsilon$. Now

$$\bigcap_{B\in\mathscr{B}(X),\,B\neq\emptyset,X}(\mathscr{F}_B\backslash\mathscr{F}^{B^c})=\{X\}$$

implies that $m(\{X\})=\varepsilon$.

EXAMPLE 3. The case $\varepsilon=1$ in the previous example models the concept of *total ignorance*. The corresponding belief function Bel_0 takes the value 0 for all proper subsets of X. Its b.p.a. is completely described by the equation $m(\{X\})=1$.

EXAMPLE 4. Let X be finite and discrete (the case treated by Shafer [9]). Then the power set is finite, too, and all subsets are closed. Writing $m(A)$ for simplicity instead of $m(\{A\})$, we get

$$\mathrm{Bel}(A)=\sum_{B\subset A} m(B).$$

Choosing $A=\varnothing$ and $A=X$ yields

$$m(\varnothing)=0 \qquad \text{and} \qquad \sum_{B\subset X} m(B)=1.$$

The plausibility function can also be expressed in terms of m by

$$\mathrm{Pl}(A)=\sum_{B\cap A\neq\varnothing} m(B).$$

3. Binary Operations on Belief Functions

Let us first recall Dempster's [3] combination of belief functions in a slightly generalized version:

Definition 3.1. Let X be a finite or countable set and m_1 and m_2 be b.p.a.'s which are both concentrated on $\mathscr{E}=\{A\subset X\colon |A|<\infty\}$ and generate belief functions Bel_1 and Bel_2, respectively. If

$$K^{-1}=\sum_{B,C\in\mathscr{E},\,B\cap C\neq\varnothing} m_1(B)\cdot m_2(C)\neq 0,$$

then the belief function generated by the b.p.a.

$$m(A)=K\sum_{B,C\in\mathscr{E},\,B\cap C=A} m_1(B)\cdot m_2(C) \tag{3.1}$$

is called the *orthogonal sum* $\mathrm{Bel}_1\oplus\mathrm{Bel}_2$. If $K^{-1}=0$, the orthogonal sum is not defined.

We notice that the sums in this Definition contain at most countably many terms so m is well defined.

Next we shall introduce a further binary operation which does not seem to be known in the literature, although its concept is very simple.

Definition 3.2. Let X be a LCS space and Bel_1, Bel_2 be belief functions on $(X,\mathscr{B}(X))$, then we call the set function Bel defined by $\mathrm{Bel}(A)=\mathrm{Bel}_1(A)\cdot\mathrm{Bel}_2(A)$ the *multiplication of* Bel_1 *and* Bel_2.

Theorem 3.3. *The multiplication of two belief functions* Bel_1, Bel_2 *on a LCS space X is again a belief function on X.*

We need the following:

Reduction Technique. Let $\sigma(A_1,\ldots,A_n)$ be the σ-algebra generated by sets $A_1,\ldots,A_n\in\mathscr{B}(X)$. For $J\subset\{1,\ldots,n\}$ the set $A(J)=\bigcap_{j\in J}A_j\cap\bigcap_{j\notin J}A_j^c$ is either empty or an atom of $\sigma(A_1,\dots, A_n)$ and $\dot\bigcup_{J\subset\{1,\dots,n\}} A(J)=X$ ($\dot\bigcup$ signifies disjoint union). The mapping

$$\operatorname{Hom}\colon \mathscr{P}(\mathscr{P}(\{1,\dots, n\}))\to\sigma(A_1,\dots, A_n)$$

$$\mathscr{J}\mapsto \dot\bigcup_{J\in\mathscr{J}} A(J)$$

is a boolean homomorphism. For $\mathscr{J}\subset\mathscr{P}(\{1,\dots, n\})$ we introduce the notation

$$\mathscr{F}_{\mathscr{J}}=\bigcap_{J\in\mathscr{J}}\mathscr{F}_{A(J)}\cap\bigcap_{J\notin\mathscr{J}}\mathscr{F}^{A(J)}$$

$$\mathscr{F}=\dot\bigcup_{\mathscr{J}\subset\mathscr{P}(\{1,\dots,n\})}\mathscr{F}_{\mathscr{J}}\qquad\text{is a partition of } \mathscr{F}.$$

Each closed set F is contained in exactly one of the $\mathscr{F}_{\mathscr{J}}$ whose index is uniquely determined by $\mathscr{J}=\{J\subset\{1,\dots, n\}\colon F\cap A(J)\neq\emptyset\}$. Now represent each non-empty set $A(J)$ by a single point $y(J)$. If m is a b.p.a. on $\mathscr{F}(X)$, it can be transferred to $Y=\{y(J)\colon A(J)\neq\emptyset\}$ by $m(\{y(J)\colon J\in\mathscr{J}\})=m(\mathscr{F}_{\mathscr{J}})$.

Proof of Theorem 3.3. The properties (i′)–(iv′) of Bel_1 and Bel_2 transfer immediately to $\mathrm{Bel}_1\cdot\mathrm{Bel}_2$. Condition (v′) is proved by utilizing the previous reduction technique:

Let m_1 and m_2 be the b.p.a.'s of Bel_1 and Bel_2, respectively, and define a probability measure on the atoms $\mathscr{F}_{\mathscr{J}}$ of $\sigma(\mathscr{F}_{\mathscr{J}})_{\mathscr{J}\subset\mathscr{P}(\{1,\dots,n\})})$ by

$$m(\mathscr{F}_{\mathscr{J}})=\sum_{\mathscr{J}',\mathscr{J}''\subset\mathscr{P}(\{1,\dots,n\}),\mathscr{J}'\cup\mathscr{J}''=\mathscr{J}} m_1(\mathscr{F}_{\mathscr{J}'})\,m_2(\mathscr{F}_{\mathscr{J}''}).\tag{3.2}$$

We show that this is the b.p.a. of the multiplication Bel of Bel_1 and Bel_2, restricted on $\sigma(A_1,\dots, A_n)$. Let $A=\dot\bigcup_{J\in\mathscr{J}} A(J)$ for some $\mathscr{J}\subset\mathscr{P}(\{1,\dots, n\})$. Then

$$\mathscr{F}^{A^c}=\dot\bigcup_{\tilde{\mathscr{J}}\subset\mathscr{J}}\mathscr{F}_{\tilde{\mathscr{J}}},$$

and we get

$$\begin{aligned}\mathrm{Bel}(A)&=\mathrm{Bel}_1(A)\,\mathrm{Bel}_2(A)=m_1(\mathscr{F}^{A^c})\,m_2(\mathscr{F}^{A^c})\\&=\Big(\sum_{\mathscr{J}'\subset\mathscr{J}} m_1(\mathscr{F}_{\mathscr{J}'})\Big)\Big(\sum_{\mathscr{J}''\subset\mathscr{J}} m_2(\mathscr{F}_{\mathscr{J}''})\Big)\\&=\sum_{\mathscr{J}'''\subset\mathscr{J}}\ \sum_{\mathscr{J}'\cup\mathscr{J}''=\mathscr{J}'''} m_1(\mathscr{F}_{\mathscr{J}'})\,m_2(\mathscr{F}_{\mathscr{J}''})\\&=\sum_{\mathscr{J}'''\subset\mathscr{J}} m(\mathscr{F}_{\mathscr{J}'''})=m(\mathscr{F}^{A^c}).\end{aligned}$$

This representation of Bel by the b.p.a. m entails that Bel is monotone of order ∞ on $\sigma(A_1, ..., A_n)$. In particular, (2.2) is satisfied.

This theorem contains an interesting special case: The powers P^k of a probability measure P (k a positive integer) are belief functions. If X is countable, (3.2) may formally be obtained from (3.1) by omitting K^{-1} and replacing intersection by union in the index of summation. Using Eq. (3.2) in this form, it is shown by induction that the b.p.a. of P^k is concentrated on those subsets of X which contain at most k elements.

Combining an arbitrary belief function Bel with the belief function of total ignorance Bel_0 yields

$$\mathrm{Bel} \oplus \mathrm{Bel}_0 = \mathrm{Bel} \qquad \text{and} \qquad \mathrm{Bel} \cdot \mathrm{Bel}_0 = \mathrm{Bel}_0.$$

Thus Bel_0 is the neutral element of the orthogonal sum. The plausibility function for total ignorance could serve as the neutral element of multiplication, but we do not use it here, because the union of belief and plausibility functions is not closed under multiplication.

4. A Measure of Non-additivity Based on an Integral

Let X be a LCS space, μ an inner regular measure, and f a μ-integrable function on X.

PROPOSITION 4.1. *The function given by* $F \mapsto \int_F f\, d\mu$ *is measurable on* $(\mathscr{F}(X), \mathscr{S}(\mathscr{F}))$.

Proof. First consider $f = 1_B$, i.e., $\int_F f\, d\mu = \mu(B \cap F)$. We have to show that $\mathscr{A}(B, c) = \{F \in \mathscr{F} : \mu(F \cap B) < c\} \in \mathscr{S}(\mathscr{F})$ for all $c > 0$. Let us prove the equality

$$\mathscr{A}(B, c) = \bigcup_{K \in \mathscr{K}(B, c)} \mathscr{F}^K$$

instead where $\mathscr{K}(B, c) = \{K \in \mathscr{K} : K \subset B$ and $\mu(K) > \mu(B) - c\}$. If $F \in \mathscr{A}(B, c)$ and $\varepsilon = c - \mu(F \cap B)$, there is a compact set $K \subset F^c \cap B$ satisfying $\mu(K) > \mu(F^c \cap B) - \varepsilon$ by inner regularity of μ. This implies $\mu(K) > \mu(B) - \mu(F \cap B) - \varepsilon = \mu(B) - c$, so $K \in \mathscr{K}(B, c)$ and $F \in \mathscr{F}^K$. For the inverse inclusion let $F \in \mathscr{F}^K$ for some $K \in \mathscr{K}(B, c)$. Then $\mu(B \cap F) \leqslant \mu(B \backslash K) < c$, so $F \in \mathscr{A}(B, c)$ and the proposition is proved for indicator functions. The extension to simple functions, nonnegative measurable functions, and integrable functions is carried out by standard methods.

Proposition 4.1 enables us to define an integral with respect to a b.p.a. m on $\mathscr{F}$:

$$\int f \otimes_\mu m = \int_{\mathscr{F}} \left(\int_F f \, d\mu \right) dm(F). \tag{4.1}$$

This integral exists since the "integrand" $\int_F f \, d\mu$ is bounded. Some properties follow immediately:

If α, β are real numbers and $f, g \in \mathscr{L}_1(\mu)$, then

$$\int (\alpha f + \beta g) \otimes_\mu m = \alpha \int f \otimes_\mu m + \beta \int g \otimes_\mu m$$

$$f \leqslant g \ \mu\text{-a.e.} \Rightarrow \int f \otimes_\mu m \leqslant \int g \otimes_\mu m$$

$$f_n \to f \text{ in } \mathscr{L}_1\text{-norm} \Rightarrow \int f_n \otimes_\mu m \to \int f \otimes_\mu m.$$

The last of these properties follows, since $\int_A |f_n - f| \, d\mu$ converges uniformly with respect to A.

Now (4.1) defines a linear, continuous functional on $\mathscr{L}_1(\mu)$, and it is well known that there is a function $h \in \mathscr{L}_\infty(\mu)$ such that

$$\int f \otimes_\mu m = \int f \cdot h \, d\mu \qquad \text{for all} \quad f \in \mathscr{L}_1(\mu).$$

Next we show how h depends on m. We write $\mathrm{Pl}(x)$ instead of $\mathrm{Pl}(\{x\})$ for simplicity.

PROPOSITION 4.2. *Suppose* Pl *is the plausibility function induced by* m. *Then*

$$\int f \otimes_\mu m = \int (f \cdot \mathrm{Pl})(x) \, d\mu(x).$$

Proof. The 1. Metrisation Theorem by Urysohn yields a metric d on X which is compatible with the topology. In addition let $\{x_1, x_2, ...\}$ be a countable and dense set in X.

First, we prove that $\mathscr{D} = \{(x, F) \in X \times \mathscr{F} : x \in F\}$ is measurable with respect to the product-σ-algebra $\mathscr{B}(X) \otimes \mathscr{S}(\mathscr{F})$. Let $U_n(x) = \{y \in X : d(x, y) < 1/n\}$. If $(x, F) \in \mathscr{D}$, then for each n there is some number k such that $x \in U_n(x_k)$ and $F \cap U_n(x_k) \neq \varnothing$. This verifies one inclusion of the equality

$$\mathscr{D} = \bigcap_{n=1}^{\infty} \left(\bigcup_{k=1}^{\infty} \left(U_n(x_k) \times \mathscr{F}_{U_n(x_k)} \right) \right). \tag{4.2}$$

To show the other inclusion we suppose $(x, F) \notin \mathscr{D}$. Choose some n such that $1/n < d(x, F)/2$. Then for all $k \geqslant 1$ at least one of the conditions $U_n(x_k) \cap F = \varnothing$ and $x \notin U_n(x_k)$ is satisfied, so (x, F) is not contained in the right-hand side of (4.2). Now Eq. (4.2) implies measurability of $\mathscr{D}$ which entails $\mathscr{B}(X) \otimes \mathscr{S}(\mathscr{F})$-measurability of the function $(x, F) \mapsto 1_F(x) \cdot f(x)$. Using Fubini's theorem we get

$$\begin{aligned}\int f \otimes_\mu m &= \int_{\mathscr{F}} \left(\int_X 1_F(x) \cdot f(x)\, d\mu(x) \right) dm(F) \\ &= \int_X f(x) \left(\int_{\mathscr{F}} 1_F(x)\, dm(F) \right) d\mu(x) \\ &= \int_X f(x) \cdot \mathrm{Pl}(x)\, d\mu(x),\end{aligned}$$

since the inner integral is equal to $m(\mathscr{F}_{\{x\}})$.

In the remaining part of this section we are going to demonstrate how this integral may be used to measure non-additivity of the set functions associated with a b.p.a. For this purpose we use the special integrand $1 = 1_X$.

1. Let us consider the case $|X| = n$; P be the uniform distribution. Then applying Proposition 4.2 yields two extreme cases, namely,

$$\int 1 \otimes_P m = \begin{cases} 1/n & \text{if } m \text{ represents a probability measure} \\ 1 & \text{if } m \text{ represents total ignorance.} \end{cases}$$

Obviously all other b.p.a.'s yield intermediate values depending on $\sum \mathrm{Pl}(x_i)$.

2. Suppose X is LCS, Bel_1 and Bel_2 are belief functions on X with b.p.a.'s m_1 and m_2, respectively, and m is the b.p.a. of $\mathrm{Bel}_1 \cdot \mathrm{Bel}_2$. Then for any measure μ and any non-negative measurable function h we get

$$\int h \otimes_\mu m_i \leqslant \int h \otimes_\mu m \leqslant \int h \otimes_\mu m_1 + \int h \otimes_\mu m_2 \qquad (i = 1, 2).$$

This is easily proved by applying Proposition 4.2 to the conjugate plausibility functions. Equality holds on the left-hand side if, e.g., Bel_1 represents total ignorance, and on the right-hand side if, e.g., $m_1(A) = m_2(A^c) = 1$ for some proper subset A.

3. Suppose X is a finite or countable set, and Bel_1, Bel_2 are belief functions whose conjugate plausibility functions will be denoted by Pl_1 and Pl_2, respectively. We assume that $\mathrm{Bel}_1 \oplus \mathrm{Bel}_2$ exists and Pl is its conjugate plausibility. Shafer's theory [9, p. 44 and Theorem 3.3] implies that

$\mathrm{Pl}(x) = K \cdot \mathrm{Pl}_1(x) \cdot \mathrm{Pl}_2(x)$, where K is given in Definition 3.1. If $K = 1$ and m_1, m_2, and m denote the b.p.a's of Bel_1, Bel_2, and $\mathrm{Bel}_1 \oplus \mathrm{Bel}_2$, respectively, this entails

$$\int h \otimes_\mu m \leqslant \int h \otimes_\mu m_i,$$

$i = 1, 2$, μ an arbitrary measure, and h an arbitrary non-negative function. So in this special case the orthogonal sum decreases the measure of non-additivity whereas the multiplication always yields an increase. It seems to us that Dempster's [3] orthogonal sum behaves more reasonably at least in this special case.

The condition $K = 1$ is satisfied, for instance, in the following situation: Let $\mathcal{A}_1$, $\mathcal{A}_2$ be independent σ-algebras on X with respect to some probability measure P. Assume that m_i is a b.p.a. concentrated on $\mathcal{A}_i$, and that $P(A) > 0$ if $m_i(A) > 0$, where $i = 1, 2$. Then $m_1(B) \cdot m_2(C) > 0$ implies $B \cap C \neq \varnothing$. Thus $K = 1$.

4. Shafer's [9] definition of discounting a belief function Bel at a rate $\varepsilon \in [0, 1]$ can be immediately extended to a LCS space:

$$\begin{aligned} &\mathrm{Bel}^\varepsilon(A) = (1 - \varepsilon)\,\mathrm{Bel}(A) \qquad \text{if} \quad A \neq X \\ &\mathrm{Bel}^\varepsilon(X) = 1. \end{aligned}$$

The conjugate equations read

$$\begin{aligned} &\mathrm{Pl}^\varepsilon(A) = (1 - \varepsilon)\,\mathrm{Pl}(A) + \varepsilon \qquad \text{if} \quad A \neq \varnothing \\ &\mathrm{Pl}^\varepsilon(\varnothing) = 0. \end{aligned}$$

This generalizes the ε-contamination which was introduced in Example 2 of Section 2. If m and m^ε denote the corresponding b.p.a.'s, and P is an arbitrary probability measure, then

$$\int 1 \otimes_P m^\varepsilon = (1 - \varepsilon) \int 1 \otimes_P m + \varepsilon \geqslant \int 1 \otimes_P m.$$

This is in accordance with our intuition that Bel^ε is "less additive" than Bel.

5. Suppose $\lambda \in (-1, \infty)$ and g_λ is a set function on $X = \{x_1, ..., x_n\}$ satisfying $g_\lambda(\varnothing) = 0$, $g_\lambda(X) = 1$, and

$$A \cap B = \varnothing \Rightarrow g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B) + \lambda g_\lambda(A)\, g_\lambda(B). \tag{4.3}$$

Then g_λ is called a *λ-additive measure* (Sugeno [11]).

The parameter λ clearly indicates the deviation of g_λ from additivity. Banon [1] and Dubois and Prade [4] have proved that g_λ is a belief function iff $\lambda \geqslant 0$ and that g_λ is a plausibility function iff $\lambda \leqslant 0$. Let $g_i = g_\lambda(\{x_i\}) < 1$, $i = 1, ..., n$. If we apply (4.3) successively, we obtain

$$g_\lambda(X) = \left(\prod_{i=1}^{n} (1 + \lambda g_i) - 1 \right) \Big/ \lambda = 1 \tag{4.4}$$

(cf. [11]). The right-hand side implies

$$\prod_{i=1}^{n} (1 + \lambda g_i) = 1 + \lambda. \tag{4.5}$$

If $g_1, ..., g_n$ are fixed and λ is replaced by $\mu - 1$, then this product is a polynomial in μ with non-negative coefficients. Thus it is convex for $\mu > 0$ or $\lambda > -1$, respectively, and (4.5) has at most two solutions. The trivial solution $\lambda = 0$ is inadmissible for (4.4). The derivative of the product in (4.5) at $\lambda = 0$ is equal to $\sum_{i=1}^{n} g_i$. So there is exactly one positive solution of (4.4) and (4.5) if $\sum g_i < 1$, and exactly one solution $\lambda \in (-1, 0)$ if $\sum g_i > 1$.

Let us investigate the relation between the parameter λ and our integral: First, suppose that $\sum g_i > 1$, i.e., $\lambda < 0$. Define another λ-additive measure by $\tilde{g}_1 > g_1$ and $\tilde{g}_i = g_i$ for $i > 1$. Then $1 + \lambda = \prod (1 + \lambda g_i) > \prod (1 + \lambda \tilde{g}_i)$ and it follows from the properties of the product that the parameter $\tilde{\lambda}$ belonging to $\tilde{g}_1, ..., \tilde{g}_n$ is smaller than λ, thus being further away from the additive case ($\lambda = 0$). Our integral reflects this simply by

$$\int 1 \otimes_\mu m(g_\lambda) = \sum_{i=1}^{n} \mu_i g_i < \sum_{i=1}^{n} \mu_i \tilde{g}_i = \int 1 \otimes_\mu m(g_{\tilde{\lambda}}),$$

where $(\mu_i)_i = (\mu(\{x_i\}))_i$ is a measure on X and $m(g_\lambda)$, $m(g_{\tilde{\lambda}})$ denote the b.p.a.'s. Now let $\sum g_i < 1$, i.e., $\lambda > 0$. A similar argument shows that because of (4.5) λ has to be increased, if one of the g_i decreases.

Applying Proposition 4.2 and (4.3) yields

$$\int 1 \otimes_\mu m(g_\lambda) = \sum_{i=1}^{n} (1 - g_\lambda(X \setminus \{x_i\})) \cdot \mu_i = \sum_{i=1}^{n} \frac{(1+\lambda) g_i}{1 + \lambda g_i} \mu_i.$$

Unfortunately, this integral may misbehave sometimes, as is demonstrated by the following

EXAMPLE. Let $n = 3$, $\mu_i = 1/3$ ($i = 1, 2, 3$), $g_2 = 0.2$, $g_3 = 0.3$. Changing the values of λ and g_1 we get the following results:

λ	g_1	$\int 1 \otimes_\mu m(g_\lambda)$
$8\frac{1}{3}$	0	0.5
7	0.0108	0.5070
3	0.1053	0.4839

However, if all g_i are identical and less than $1/n$, there is a monotone relation between λ and $\int 1 \otimes_\mu m(g_\lambda)$:

Let $\bar{g} = g_1 = \cdots = g_n$, then $\bar{g} = ((1+\lambda)^{1/n} - 1)/\lambda$ and

$$\int 1 \otimes_\mu m(g_\lambda) = \frac{(1+\lambda)\bar{g}}{1+\lambda\bar{g}} \sum_{i=1}^{n} \mu_i = \frac{(1+\lambda)((1+\lambda)^{1/n} - 1)}{(1+\lambda)^{1/n}} \sum_{i=1}^{n} \mu_i.$$

Using Bernoulli's inequality a straightforward calculation shows that the derivative of the term on the right-hand side is positive.

5. A Generalization of the Lebesgue Integral

For $f \in \mathscr{L}_1(\mu)$ and $\mu_F(B) = \mu(B \mid F) = \mu(B \cap F)/\mu(F)$ the conditional measure if $\mu(F) > 0$, let us have a look at the following preliminary definition of a new integral

$$\oint f \otimes_\mu m = \int_{\mathscr{F}} \left(\int_F f \, d\mu_F \right) dm(F). \tag{5.1}$$

This requires some discussion.

First, μ_F is not defined if $\mu(F) = 0$. Therefore we should eliminate the set $\mathscr{F}_0 = \{F \in \mathscr{F} : \mu(F) = 0\}$ from the range of integration of the outer integral. By Proposition 4.1, $\mathscr{F}_0$ is a measurable subset of $\mathscr{F}$. In order to avoid a loss of relevant information, we have to presuppose that $m(\mathscr{F}_0) = 0$. Then Fubini's theorem may be applied to

$$\int_{\mathscr{F} \setminus \mathscr{F}_0} \left(\int_X 1_F(x) \cdot f(x)/\mu(F) \, d\mu(x) \right) dm(F),$$

since the integrand is $\mathscr{B}(X) \otimes \mathscr{S}(\mathscr{F})$-measurable. But even for such a reasonable example as the ε-contamination of a continuous probability measure $m(\mathscr{F}_0) > 0$.

Also the argument utilizing the equivalence of the dual space $\mathscr{L}_1^*(\mu)$ and the space of the a.s. bounded functions on X cannot be used in general. This is due to the fact that the functional on $\mathscr{L}_1(\mu)$ defined by (5.1) is not always continuous: It may happen that $\int |f_n| \, d\mu$ converges to 0, but for some decreasing sequence $\{F_n\}$, $\int_{F_n} f_n \, d\mu_{F_n}$ increases so quickly and $m(\{F \in \mathscr{F} : F \subset F_n\})$ converges so slowly that $\oint f_n \otimes_\mu m$ does not converge to 0.

In order to avoid these difficulties, we restrict the definition of the integral to the following case:

DEFINITION 5.1. Suppose X is a finite or countable set, m is a b.p.a. concentrated on $\mathscr{E}$ (cf. Definition 3.1), μ is a measure which is positive on non-empty subsets of X, and f is a function on X. Then the *integral of f* with respect to m (or the associated belief or plausibility function) is defined by

$$\oint f \otimes_\mu m = \int_{\mathscr{E}} \left(\int_A f \, d\mu_A \right) dm(A).$$

If the b.p.a.'s of two belief functions are restricted to $\mathscr{E}$, (3.1) and the remark following Theorem 3.3 reveal that this property is satisfied by the orthogonal sum and the multiplication, too.

With regard to the condition concerning μ, one should adopt an absolute continuity of the type $\mu(A) = 0 \Rightarrow \mathrm{Pl}(A) = 0$. Then, sets of μ-measure 0 can be removed from X and μ becomes positive.

Again, some properties of the integral follow immediately:

Suppose α, β are real numbers and f, g are bounded functions on X, then

$$\oint (\alpha f + \beta g) \otimes_\mu m = \alpha \oint f \otimes_\mu m + \beta \oint g \otimes_\mu m$$

$$f \leqslant g \Rightarrow \oint f \otimes_\mu m \leqslant \oint g \otimes_\mu m$$

$$\oint \alpha \otimes_\mu m = \alpha.$$

If $\{f_n\}$ is a pointwise non-decreasing sequence of bounded functions, then

$$\sup_n \oint f_n \otimes_\mu m = \oint (\sup_n f_n) \otimes_\mu m.$$

This is proved by applying the classical theorem of monotone convergence to the functions $h_n(A) = (\int_A f_n \, d\mu)/\mu(A)$.

A function f is called *m-integrable* if $\oint |f| \otimes_\mu m < \infty$. If $f_n \to f$ pointwise and there is a function $\bar{f}$, which is m-integrable and an upper bound for all $|f_n|$, then

$$\oint f_n \otimes_\mu m \to \oint f \otimes_\mu m.$$

In order to prove this again let $h_n(A) = (\int_A f_n \, d\mu)/\mu(A)$ and define $\bar{h}(A)$ and

$h(A)$ analogously. Then $h_n(A) \to h(A)$ and $|h_n(A)| \leqslant \bar{h}(A)$ for all n and A. Now apply Lebesgue's bounded convergence theorem.

PROPOSITION 5.2. *If a b.p.a. m represents a probability measure, the integral defined in Definition* 5.1 *is the Lebesgue integral.*

Proof. The b.p.a. m of a probability measure P is restricted to singletons, so

$$\oint f \otimes_\mu m = \sum_{x \in X} \left(\int_{\{x\}} f \, d\mu_{\{x\}} \right) m(\{x\}) = \sum_{x \in X} f(x) \, P(\{x\})$$
$$= \int f \, dP.$$

Since the integral $\oint f \otimes_\mu m$ is linear and satisfies the theorem of monotone convergence, a probability measure is defined by

$$\Pr(A) = \oint 1_A \otimes_\mu m. \tag{5.2}$$

Writing down the appropriate series yields

$$\Pr(A) = \mathrm{Bel}(A) + \sum_{B \not\subset A} \frac{\mu(A \cap B)}{\mu(B)} m(B)$$
$$= \mathrm{Pl}(A) - \sum_{B \cap A \neq \varnothing} \frac{\mu(B \cap A^c)}{\mu(B)} m(B).$$

Hence $\mathrm{Bel}(A) \leqslant \Pr(A) \leqslant \mathrm{Pl}(A)$ for all $A \subset X$. In particular

$$\mathrm{pr}_i = \Pr(\{x_i\}) = \sum_{B \ni x_i} \frac{\mu(\{x_i\})}{\mu(B)} m(B). \tag{5.3}$$

If f is m-integrable, the bounded convergence theorem enables us to write

$$\oint f \otimes_\mu m = \sum_i f(x_i) \oint 1_{\{x_i\}} \otimes_\mu m = \sum_i f(x_i) \, \mathrm{pr}_i.$$

Thus the integral introduced here inherits its properties from the Lebesgue integral. In addition, it is easily computed, if the pr_i are calculated in advance.

The expression for pr_i shows that the probability $m(B)$ is distributed on B according to weights, which are determined by the measure μ. In the finite case $X = \{x_1, ..., x_n\}$ Dempster [3] has also constructed probability measures on X by distributing $m(B)$ amongst the elements of B, but he allows arbitrary weights for each set B. By his method any probability measure Pr satisfying $\Pr(B) \geqslant \mathrm{Bel}(B)$ can be obtained. (The inequality $\Pr(B) \leqslant \mathrm{Pl}(B)$ holds simultaneously for all B, as $1 - \Pr(B^c) \leqslant 1 - \mathrm{Bel}(B^c)$.) Each permutation π on X defines an extremal point of the convex cone of these probability measures by

$$p_i^\pi = \sum_{\substack{B \subset X \\ \pi(i) = \min\{\pi(j):\, x_j \in B\}}} m(B)$$

(Dempster [3]). Hence (p_i^π) allocates the probability $m(B)$ completely to that $x_i \in B$ which becomes the first element of B under the permutation π. If Pr is defined as in (5.2), the extremal points $(p_i^\pi)_{i=1,\dots,n}$ obviously cannot be reached, but they can be approximated asymptotically:

Proposition 5.3. *Suppose $X = \{x_1, \dots, x_n\}$. For each permutation π on X and each $\varepsilon > 0$ there is a measure μ such that $|\mathrm{pr}_i - p_i^\pi| < \varepsilon$ for all $i \leqslant n$.*

Proof. Let $\delta > 0$ and $\sum_{k=2}^n \delta^k < 1$. Define μ by

$$\mu_i = \mu(\{x_i\}) = \begin{cases} \delta^{\pi(i)} & \text{if } \pi(i) > 1 \\ 1 - \sum_{k=2}^{n} \delta^k & \text{if } \pi(i) = 1. \end{cases}$$

Then

$$\mathrm{pr}_i = \begin{cases} \delta^{\pi(i)} \sum_{B \ni x_i} m(B)/\mu(B) & \text{if } \pi(i) > 1 \\ \left(1 - \sum_{k=2}^{n} \delta^k\right) \sum_{B \ni x_i} m(B)/\mu(B) & \text{if } \pi(i) = 1. \end{cases}$$

Now, if $\delta \to 0$, straightforward computations show that the weight of $m(B)$ approaches 1 if $\pi(i) = 1$ or $\pi(i) = \min\{\pi(j): x_j \in B\}$, and that the weight approaches 0 if $\pi(i) > \min\{\pi(j): x_j \in B\}$. Since there are only finitely many terms involved, δ can be chosen sufficiently small in order to achieve $|\mathrm{pr}_i - p_i^\pi| < \varepsilon$.

Let us now demonstrate the computation of the integral for squares and cubes of probability measures P. Suppose μ is a measure, f is a bounded function, and $m^{(2)}$ is the b.p.a. of P^2, i.e., $m^{(2)}(\{x_i\}) = p_i^2$, $m^{(2)}(\{x_i, x_j\}) = 2p_i p_j$ $(i \neq j)$, and $m^{(2)}(A) = 0$ if $|A| > 2$. Then

$$\mathrm{pr}_i = p_i^2 + \mu_i \sum_{j \neq i} 2p_i p_j / (\mu_i + \mu_j)$$

and hence

$$\oint f \otimes_\mu m^{(2)} = \sum_{i=1}^{\infty} f(x_i)\, p_i \left(p_i + 2\mu_i \sum_{j \neq i} p_j/(\mu_i + \mu_j) \right).$$

If all μ_i are the same, we obviously get the Lebesgue integral $\int f\,dP$. This is *not* true if we take P^3 instead, as can be checked easily.

A Comparison with Choquet's and Sugeno's Integral

The notion of capacities (Choquet [2]) is more general than that of belief and plausibility functions. So *Choquet's integral* [2]

$$\int_0^{\sup f} T(\{f > t\})\, dt$$

for non-negative functions f and capacities T can be applied to Bel and Pl. It was pointed out by Choquet [2] that

$$\int_0^{\sup f} \mathrm{Bel}(\{f > t\})\, dt = \int (\inf_{x \in F} f(x))\, dm(F)$$

and

$$\int_0^{\sup f} \mathrm{Pl}(\{f > t\})\, dt = \int (\sup_{x \in F} f(x))\, dm(F).$$

This entails

$$\int_0^{\sup f} \mathrm{Bel}(\{f > t\})\, dt \leqslant \oint f \otimes_\mu m \leqslant \int_0^{\sup f} \mathrm{Pl}(\{f > t\})\, dt, \tag{5.4}$$

if all integrals are defined. Equality holds if f is constant on any set $B \subset X$ satisfying $m(B) > 0$; in particular, if $\mathrm{Bel} = \mathrm{Pl}$ is a probability measure, all these integrals are equal to the Lebesgue integral.

Sugeno [11] defines a *fuzzy measure* to be a set function gr on a σ-algebra which is monotone with respect to set inclusion, continuous with respect to increasing and decreasing sequences of sets, and normalized by $\mathrm{gr}(\varnothing) = 0$ and $\mathrm{gr}(X) = 1$. If $f\colon X \to [0, 1]$ is a measurable function, he defines an integral

$$⨍ f \circ \mathrm{gr} = \sup_{\alpha \in [0,1]} (\min(\alpha, \mathrm{gr}(F_\alpha))), \qquad \text{where} \quad F_\alpha = \{f \geqslant \alpha\}.$$

Sugeno [11] has proved that the difference between Lebesgue's and his own integral is not greater than $\frac{1}{4}$ if gr is σ-additive. A comparison with our own integral yields the following result:

PROPOSITION 5.4. *If the assumptions of Definition* 5.1 *hold and the range of f is in* $[0, 1]$, *then*

(i) $-\frac{1}{4} \leqslant \oint f \otimes_\mu m - ⨍ f \circ \mathrm{Bel} < 1$,

(ii) $\frac{1}{4} \geqslant \oint f \otimes_\mu m - ⨍ f \circ \mathrm{Pl} > -1$.

In general, these inequalities cannot be improved.

Proof. (i) Suppose $\alpha \in [0, 1]$ and $A \in F_\alpha$. Then $(\int_A f\, d\mu)/\mu(A) \geqslant \alpha$, and this implies

$$\begin{aligned}\oint f \otimes_\mu m &\geqslant \sum_{A \subset F_\alpha} \left(\int_A f\, d\mu\right)\Big/\mu(A) \cdot m(A)\\ &\geqslant \alpha \cdot \mathrm{Bel}(F_\alpha)\\ &\geqslant (\min(\alpha, \mathrm{Bel}(F_\alpha)))^2.\end{aligned}$$

Now $\oint f \otimes_\mu m - ⨍ f \circ \mathrm{Bel} \geqslant (⨍ f \circ \mathrm{Bel})^2 - ⨍ f \circ \mathrm{Bel} \geqslant -\frac{1}{4}$. If $\mathrm{Bel} = P$ is σ-additive, $P(A) = \frac{1}{2}$ and $f = \frac{1}{2} \cdot 1_A$, then equality holds on the left-hand side.

With regard to the inequality on the right-hand side of (i), it is easy to see that $\oint f \otimes_\mu m = 1$ and $⨍ f \circ \mathrm{Bel} = 0$ do not hold simultaneously. However, the difference in (i) can be arbitrarily close to 1:

Let $X = \{x_1, ..., x_n\}$, $\mu(\{x_i\}) = 1$ for $i = 1, ..., n$, $k < n$, $f(x_1) = \cdots = f(x_k) = 0$, $f(x_{k+1}) = \cdots = f(x_n) = 1$, and $m(\{x_k, x_{k+1}, ..., x_n\}) = 1$. Then

$$\oint f \otimes_\mu m = \frac{n-k}{n-k+1} m(\{x_k, ..., x_n\}) \to 1 \qquad \text{as} \quad n \to \infty,$$

but $⨍ f \circ \mathrm{Bel} = 0$.

(ii) Write $M = ⨍ f \circ \mathrm{Pl}$ and $F_{M+0} = \{f > M\}$, then

$$\mathrm{Pl}(F_{M+0}) \leqslant M \leqslant \mathrm{Pl}(F_M) \qquad \text{(cf. [11])}.$$

Since

$$\oint f \otimes_\mu m \leqslant \sum_{A \cap F_{M+0} \neq \varnothing} 1 \cdot m(A) + \sum_{A \cap F_{M+0} = \varnothing} M \cdot m(A),$$

we deduce

$$\begin{aligned}\oint f \otimes_\mu m - ⨍ f \circ \mathrm{Pl} &\leqslant \mathrm{Pl}(F_{M+0}) + M(1 - \mathrm{Pl}(F_{M+0})) - M\\ &\leqslant M(1 - M) \leqslant \tfrac{1}{4}.\end{aligned}$$

Equality holds if $\mathrm{Pl} = P$ is σ-additive, $P(A) = \frac{1}{2}$, and $f = \frac{1}{2} 1_A + 1_{A^c}$. Again, it is easily checked that $⨍ f \circ \mathrm{Pl} = 1$ and $\oint f \otimes_\mu m = 0$ do not hold simultaneously. The following example shows, however, that the difference in (ii) may be arbitrarily close to -1:

Suppose $X=\{x_1,\ldots,x_n\}$, $\mu(\{x_i\})=1$ for $i=1,\ldots,n$, $f(x_1)=\cdots=f(x_{n-1})=0$, $f(x_n)=1$, and $m(X)=1$. This implies $\int f\circ \mathrm{Pl}=1$ and $\oint f\otimes_\mu m=m(X)/n\to 0$ as $n\to\infty$ and hence everything is shown.

Remarks. (1) Alternatively one can fix n and modify μ in the two previous examples in order to achieve the same results.

(2) The upper bound $\frac{1}{4}$ of the difference between Lebesgue's and Sugeno's integral cannot be improved. On one side of our inequalities this bound still applies if we compare Sugeno's integral with ours (left-hand side of the inequalities above). On the other hand, these differences may be extremely large, since the integrals take only values in [0, 1].

(3) Combining the inequalities in Proposition 5.4 yields

$$\int f\circ \mathrm{Bel}-\frac{1}{4}\leqslant \oint f\otimes_\mu m\leqslant \int f\circ \mathrm{Pl}+\frac{1}{4}.$$

This is an analogous result to (5.4).

Both Choquet's and Sugeno's integrals are defined for very general classes of set functions on any measurable space. They are not linear in their integrands and have not been defined for all real, measurable functions. In this sense they are distinct from the type of integrals introduced in this paper.

Acknowledgments

I want to thank W. Bühler, H.-J. Schuh, and S. Weber for helpful discussions and valuable comments.

References

1. G. Banon, Distinction between several subsets of fuzzy measures, *Fuzzy Sets and Systems* **5** (1981), 291–305.
2. G. Choquet, Theory of capacities, *Ann. Inst. Fourier* (*Grenoble*) **5** (1953/1954), 131–295.
3. A. P. Dempster, Upper and lower probabilities induced by a multivalued mapping, *Ann. Math. Statist.* **38** (1967), 325–339.
4. D. Dubois and H. Prade, "Fuzzy Sets and Systems: Theory and Applications," Academic Press, New York, 1980.
5. P. J. Huber and V. Strassen, Minimax tests and the Neyman–Pearson Lemma for capacities, *Ann. Statist.* **1** (1973), 251–263.
6. G. Matheron, "Random Sets and Integral Geometry," Wiley, New York, 1975.
7. H. T. Nguyen, On random sets and belief functions, *J. Math. Anal. Appl.* **65** (1978), 531–542.

8. G. R. SHAFER, "Allocations of Probability: A Theory of Partial Belief," Ph.D. thesis, University of Michigan, Ann Arbor, 1973.
9. G. R. SHAFER, "A Mathematical Theory of Evidence," Princeton Univ. Press, Princeton, N.J., 1976.
10. G. R. SHAFER, Allocations of probability, *Ann. Probab.* **7** (1979), 827–839.
11. M. SUGENO, "Theory of Fuzzy Integrals and Its Application," Ph.D. thesis, Tokyo Institute of Technology, 1974.
12. P. WALLEY AND T. L. FINE, Towards a frequentist theory of upper and lower probability, *Ann. Statist.* **10** (1982), 741–761.